				The PPR Panel disagrees with these two sentences. In some studies dose and exposure duration can be quantified (e.g., smoking). Also, associations between exposures and outcomes can be established in the lack of good exposure data.
58	Ministero della Salute	ITA	2.3 Exposure	Page 14 Lines 616-619: This is one of the specific points in which the lack of toxicokinetic consideration is particularly evident. The sentence There may be differences in absorption and metabolism via different routes (dermal, inhalation and oral) is not complete and can be misleading: differences are not limited to absorption, but systemic bioavailability can be much more relevant (occurring also with the same level of absorption) when different route of exposure are compared. See the BPA story, as an example. In addition, interindividual differences due to change in kinetics parameters can determine a different internal dose among the general population. The identification of group with different susceptibility is not limited to differences in cell response (i.e. toxicodynamics) *EFSA Response:* *The following text has been included at line 618: "Pharmacokinetic differences among individuals may result in differing systemic or tissue/organ doses even where the absorbed external doses may appear similar".*
59	SYNGENTA	GBR	2.4 Health outcomes	It is common practice in many observational epidemiological investigations to rely on self-reported disease outcomes. The utility of using such metrics in assessing the association between pesticide exposure and disease is of limited use. This limitation can be best overcome by independent, blinded assessment of disease status by a medical expert assigned to a case-control study or a case-control study nested within a pre-established cohort. *EFSA Response:* *Section 2.4 (health outcomes) is placed under chapter 2, where general principles of epidemiological studies are summarized under the viewpoint of the PPR Panel. This section does not mention self-reported disease outcomes, which otherwise is mentioned elsewhere (section 3.1, line 959). Because this comment can be best considered a recommendation, it has been considered in section 8.1.c.* *The following text has been included at line 2435 as a new point c.1): "Self-reported health outcomes should be avoided or confirmed by independent, blinded assessment of disease status by a medical expert assigned to the study."* *The following sentence has been added in Line 1088: "Self-reported disease outcomes are frequently used in epidemiological research because of the difficulty of verifying responses in studies with large samples and limited funds, among other reasons. Although a number of studies have examined agreement between self-reported outcomes and medical records, the lack of verification of such metrics can lead to misclassification, particularly in large population-based studies, which may detract reliability to the associations found."*
60	Ministero della Salute	ITA	2.4 Health outcomes	Page 15 Lines 635-640. The whole paragraph is a bit off, and particularly the last sentence does not make sense. Mortality and disease registries should be better presented. Among the weaknesses there is the delay with which data become available (e.g. for many cancer registries) Page 15 Lines 646-647 The sentence starting with Harmonisation … does not make sense. Page 15 Line 679: ' evaluating whether if…' if should be deleted

				EFSA Response: *This paragraph has been added as disease registries can be used as a source of health outcomes in epidemiological studies, particularly mortality or cancer registries.* *Line 638: If the case definition varies from country to country, it is hard to pool data from different countries on the same disease in order to obtain sound associations.* *The term 'Harmonization' in line 646 means that diagnostic criteria should be harmonized to increase quality of epi studies, particularly when the information is obtained from different settings.* *The following text has been added in line 635: "Disease registries contain clinical information on patients on diagnosis, treatment, and outcome. These registries periodically update patient information and thus can provide useful data for epidemiologic research."* *For clarification, the sentence in lines 638-639 has been deleted (also, diagnoses can be recorded in refined or relatively crude format).* *The following text has been added in line 640: "Registry data present many opportunities for meaningful analysis, but the degree of data completeness and validity may challenge making appropriate inferences. Also, changes in coding conventions over the lifetime of the database may have an impact on retrospective database research."* *In line 679 the word 'if' has been deleted.*
61	SYNGENTA	GBR	2.5 Statistical analysis and reporting	No comment on this description of statistical methodology applicable to the analysis and interpretation of observational epidemiology. The reader should also be referred to standard textbooks where statistical methodology used in epidemiology studies is presented in a more comprehensive manner (e.g. Rothman et al., 2008; Thomas et al., 2009). *EFSA Response:* *See reply to comment #53.*
62	ECPA	BEL	2.5.2 Modelling exposure-health relationship	This section of the scientific opinion first introduces the use of models to characterize the relationship between exposure levels and the risk of health effects of interest. The opinion would benefit from an expanded discussion of dose-response in general, how dose-response analyses should be conducted, and how dose-response analyses can contribute to an evaluation of causation. Here it is also important to draw the distinction between exposure and internal dose (i.e., the amount of the substance that is absorbed by the body). Biomonitoring studies show why this is critical; the number of days of pesticide use or even generic exposure models that purport to assess intensity are not good surrogates for internal dose. *EFSA Response:* *This suggestion is beyond the scope of the opinion.* Line 695: "Failure to reject the null hypothesis does not necessarily mean that no association is present because the study may not have sufficient power to detect it". Although it should go without saying that a null result could indicate there is no relationship between an exposure and outcome, it would be useful to explicitly state this in the scientific opinion.

				EFSA Response: *The text already suggests that no association is a possibility.*
63	personal	USA	2.5.2 Modelling exposure-health relationship	Lines 706-718. The discussion of statistical and biological significance is an important one. It's true that they are not always the same thing. Although there is a statement that an association that is "statistically significant may be or may not be biologically relevant and vice versa", in practice, this is almost never taken to be true in regulatory policy and epidemiologic results that are statistically significant may be dismissed as "not biologically relevant", but it is almost never the case that non-statistically significant results are determined to be "biologically relevant". *EFSA Response:* *The following text has been added in line 710: "While epidemiological results that are statistically significant may be dismissed as "not biologically relevant", non-statistically significant results are seldom determined to be "biologically relevant".* The document includes several contradictory positions about what constitutes a "good" or "reliable" epidemiologic study. For example, in section 2.2, in lines 573-580, the authors suggest that it's important to have a study that represents the underlying population, and suggesting drawing a representative sample from which to draw inferences. However, on lines 802-811, the authors suggest that restricting the sample to avoid any potential confounders is the preferred method for accounting for potential confounders. These are two contradictory design considerations. In practice, there is almost always a trade-off between strengths and limitations of different decisions that must be accounted for in designing an epidemiologic study and these (and other) choices should be driven by the specific question being asked and the study population. As written, this document seems to make little effort to distinguish which of many criticisms of epidemiologic studies are most important to consider, leaving the reader to wonder if there would ever be an instance where an epidemiologic study would not be so flawed as to not be useful for the purposes of regulation. *EFSA Response:* *Thanks for spotting this. The text in line 579 has been removed (Representative samples can be achieved…). Section 2.2 acknowledges that a representative sample of the target population is a key point in epi studies. In turn, what is mentioned in lines 802-811 is that stratification of the selected population can be a useful way to control for a given confounding factor.*
64	US EPA	USA	2.5.2 Modelling exposure-health relationship	Line 685 In reality control group is not necessarily not exposed. It is the group of lowest exposure *EFSA Response:* *The text in line 683 has been slightly changed as following: "The relative risk (RR) in cohort studies estimates the relative magnitude of an association between exposure and disease comparing those who are exposed (or those that have a higher exposure level) with those that are not exposed (or those that have a lower exposure level). It indicates the likelihood of developing the disease in the exposed group relative to those who are not (or less) exposed".*

				Line 688 It is not just ALL diseased, it is diseased from a particular disease of interest, which is usually of interest and not overall disease level. This inaccuracy is something that applies throughout the document Line 689 could give an example of what statistical approaches are used for relative measure - i.e. different regressions - logistic or proportional hazards Line 690 Median change is usually examined using non-parametric statistics and not ANOVA Line 692. Normal distribution of the outcome is not necessary for using "other parametric statistics". *EFSA Response:* *The text in line 690 has been slightly changed as following: "For continuous outcome measures, mean or median change in the outcome are often examined across different level of exposure; either through analyses of variance or through other parametric statistics".* Lines 746-779. This section is totally out of place here. Effect size magnification is only relevant when only significant results are picked. Unlike other issues (confounding, effect modification, selection, analysis, etc.), this is not property of a study, but something who analyzes studies does. This discussion is not appropriate here. This needs to be discussed in conjunction of meta-analytic approaches that look at studies without looking at individual significance. *EFSA Response:* *This section has been moved to a new (last) bullet point under heading "Interpretation of statistically significant difference" (section 2.5.2.).*
65	Ministero della Salute	ITA	2.5.2 Modelling exposure-health relationship	Page 15 line 673. Chapter title: "exposure-outcome" or "exposure-health outcome", relationship, instead of "exposure-health" is suggested *EFSA Response:* *Addressed.* Page 16, lines 693-694: This sentence does not make sense. It needs rewriting. Page 16, lines 695-827. Interpretation of results should not be under the chapter modelling. It should have a chapter by itself. Moreover, interpretation should occur after the steps in the following chapters (validity, sensitivity analysis etc have been considered), and thus should be moved after those. *EFSA Response:* *The PPR Panel considers that no changes are needed in page 16.* Page 17-18: when confounders are described, a more detailed description on their identification could be beneficial (e.g. see the general comment on kinetics and interactions among chemicals). Page 18 lines 832-835: when genetic polymorphism is mentioned it should be clear that a biological plausibility should be included in the reporting of results: to include genotyping of the population for some allelic variants without a biological basis is most of the times waste of time and resources

				Genetic polymorphisms are mentioned only s effect modifiers *EFSA Response:* *The following text has been included in line 793: "For instance, because agriculture exposures cover many different exposure categories, farmers are likely to be more highly exposed than the general population to a wide array of risk factors, including biological agents (soil organisms, livestock, farm animals), pollen, dust, sunlight and ozone amongst others, which may act as potential confounding factors".* *Genetic polymorphisms are also mentioned in chapter 3.3. (relevance of study populations).*
66	ECPA	BEL	2.6 Study validity	Lines 881-882: the scientific opinion states that selection bias can "generally be dealt with by careful design and conduct of a study". This section should reference the other sections of the opinion that provide detailed information on how a study can avoid this bias. *EFSA Response:* *Some cross-references to chapters 4, 6 and 8 have been included under chapter 2.6.*
67	Centre F Baclesse	FRA	2.6 Study validity	916-927: This section should indicate what situations must occur for non-differential misclassification to bias relative risks away from the null. For epidemiologic studies that can evaluate an exposure-response gradient, the Panel should indicate how non-differential misclassification can bias the relative risks upward for the highest exposure category. This is important because the upper category is critical for the slope. Again rather than just saying it might happen, the Panel provide documentation of situations where it has. *EFSA Response:* *See comments #49 and 59.*
68	Université de Bordeaux	FRA	2.6 Study validity	916-927: This section should indicate what situations must occur for non-differential misclassification to bias relative risks away from the null. For epidemiologic studies that can evaluate an exposure-response gradient, the Panel should indicate how non-differential misclassification can bias the relative risks upward for the highest exposure category. This is important because the upper category is critical for the slope. Again rather than just saying it might happen, the Panel provide documentation of situations where it has. *EFSA Response:* *Same text as comment #67.*
69	SYNGENTA	GBR	2.6 Study validity	Study validity is more than just the absence of bias. Study validity is not ensured by the use of methods to assess the temporal relationship between dose and response or the use of sensitivity analyses. The generalizability of the result from the population under study to a broader population should also be included in any discussion on validity. In addition, it should be acknowledged that the results from a single study take on increased validity when it is replicated in independent investigations conducted on other populations of individuals at risk of developing the disease.